5/14/18

"What goes up doesnt always have to come down." -Unknown

HW: "Intro to Statistics" homework section

Test 2 on Wednesday 5/30

AIM: What are Statistics?

Warm Up:

1. The graph of $y = 3\sin(3x - \frac{\pi}{3}) - 4$ is a transformation of $y = 3\sin(3x)$. Describe the transformation completely.

Shift right $\frac{\pi}{3}$ units

down 4 units

SINUSOIDAL MODELING
COMMON CORE ALGEBRA II HOMEWORK

APPLICATIONS

1. A ball is attached to a spring, which is stretched and then let go. The height of the ball is given by the sinusoidal equation $y = -3.5\cos\left(\frac{4\pi}{5}t\right) + 5$, where y is the height above the ground in feet and t is the number of seconds since the ball was released.

(a) At what height was the ball released at? Show the calculation that leads to your answer.

$$\boxed{y_{\min} = 5 - 3.5 = 1.5 \text{ ft}}$$

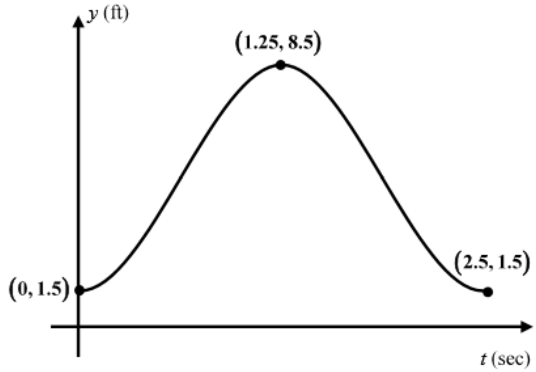
(b) What is the maximum height the ball reaches?

$$\boxed{y_{\max} = 5 + 3.5 = 8.5 \text{ ft}}$$

(c) How many seconds does it take the ball to return to its original position?

$$\boxed{\frac{4\pi}{5}P = 2\pi \Rightarrow P = 2\pi \cdot \frac{5}{4\pi} = \cancel{2\pi} \cdot \frac{5}{2 \cdot \cancel{2\pi}} = \frac{5}{2} = 2.5 \text{ sec}}$$

(d) Draw a rough sketch of one complete period of this curve below. Label maximum and minimum points.



2. An athlete was having her blood pressure monitored during a workout. Doctors found that her maximum blood pressure, known as systolic, was 110 and her minimum blood pressure, known as diastolic, was 70. If each heartbeat cycle takes 0.75 seconds, then determine a sinusoidal model, in the form $y = A\sin(Bt) + C$, for her blood pressure as a function of time t in seconds. Show the calculations that lead to your answer.

$$C = \frac{70 + 110}{2} = \frac{180}{2} = 90$$

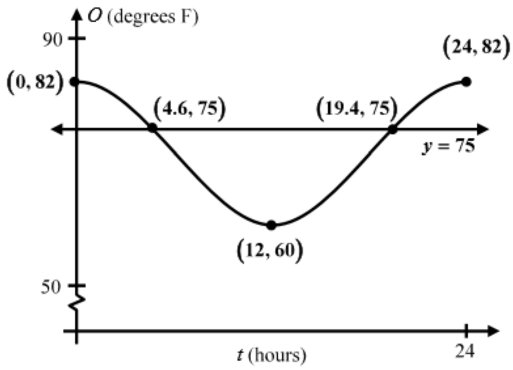
$$A = \frac{110 - 70}{2} = \frac{40}{2} = 20$$

$$0.75B = 2\pi \Rightarrow B = \frac{2\pi}{0.75} = \frac{2}{0.75}\pi = \frac{8}{3}\pi$$

$$\Rightarrow \quad \boxed{y = 20\sin\left(\frac{8\pi}{3}t\right) + 90}$$

3. On a standard summer day in upstate New York, the temperature outside can be modeled using the sinusoidal equation $O(t)=11\cos\left(\frac{\pi}{12}t\right)+71$, where t represents the number of hours since the peak temperature for the day.

(a) Sketch a graph of this function on the axes below for one day.



(b) For $0 \le t \le 24$, graphically determine all points in time when the outside temperature is equal to 75 degrees. Round your answers to the nearest tenth of an hour.

According to our graph at the left the temperature outside is 75 °F at: $t = 4.6$ and 19.4 hours

4. The percentage of the moon's surface that is visible to a person standing on the Earth varies with the time since the moon was full. The moon passes through a full cycle in 28 days, from full moon to full moon. The maximum percentage of the moon's surface that is visible is 50%. Determine an equation, in the form $P = A\cos(Bt)+C$ for the percentage of the surface that is visible, *P*, as a function of the number of days, *t*, since the moon was full. Show the work that leads to the values of *A*, *B*, and *C*.

$$C=\frac{0+50}{2}=\frac{50}{2}=25$$

$$A=\frac{50-0}{2}=\frac{50}{2}=25$$

$$28B=2\pi \Rightarrow B=\frac{2\pi}{28}=\frac{\pi}{14}$$

$$\Rightarrow \quad \boxed{y=25\cos\left(\frac{\pi}{14}t\right)+25}$$

5. Evie is on a swing thinking about trigonometry (no seriously!). She realizes that her height above the ground is a periodic function of time that can be modeled using $h=3\cos\left(\frac{\pi}{2}t\right)+5$, where t represents time in seconds. Which of the following is the range of Evie's heights?

(1) $2 \le h \le 8$ (3) $3 \le h \le 5$

(2) $4 \le h \le 8$ (4) $2 \le h \le 5$

$h_{min} = 5-3 = 2$ ft

$h_{max} = 5+3 = 8$ ft

(1)

Data is everywhere. It's in our newspapers, it's in our science classes, it shows up in economics, medicine and anywhere else that **variability** occurs. Variability is simply the property of **outcomes** being different. The tools of statistics are designed to **explain this variability**.

There are many types of variability. It is good to understand these sources in order to minimize the ones that we are not studying.

(a) **Observational or Measurement Variability**: Variability that is introduced due to either our measuring instruments not being precise enough or differences in how two different people read the measurement.

Reaction time of 2 people
timing a race.

(b) **Natural Variability or Inter-Individual Variability**: Variability that accounts for the fact that members of a populations are simply different.

Two people eat the same thing
but gain different amounts of weight.

(c) **Induced Variability**: This type of variability is in marked contrast to natural. It occurs because we have assigned our population or sample to two or more **treatment** groups and then observe the variability between the groups.

Giving one group medicine and giving
another group a placebo.

(d) **Sample Variability**: This is the type of variability that occurs when we take multiple **samples** from a **population** randomly. These samples will be different due to the randomness of the sampling process.

Choose 10 people
some may get 7 boys 3 girls
another may get 5 boys 5 girls

Remember, through all of our work in this unit, we are really trying to explain the variability of data within either a population or a sample and then using this to determine if the variability can be attributed to one of the factors above to the exclusion of the others.

There are many different situations in which we collect data. They have important differences and all of them depend on **randomization** in one way or another.

The three major types of ways to collect data are described below. Let's give an example of each and explain how **randomization** is part of each method. Randomization is used primarily to eliminate variability caused by some type of **bias**. ← not fair

(a) **Surveys**: Collections of data from a population where variability is not induced by treatments but by the sample itself (sampling variability).

Polls, phone surveys, email survey, questionnaire

(b) **Observational Studies**: Collections of data from a population where assignment of individuals from the population into **treatment groups** is **not** under the control of those performing the study.

Look at whats happening.

(c) **Experimental Studies**: In experimental studies individuals are assigned randomly to treatment groups in order to determine the effect of the treatment on the variability of the data. In these cases, the assignment, although random, is under the control of those performing the study.

control group - doesnt get it.

Test group - gets medicine

Random sampling is critical for being able to minimize variability due to **sampling bias**. Random sampling can be done using a variety of different techniques. Simple random sampling can be accomplished using a random number table.

***Exercise* #1:** A list of 10 people's heights, in inches, is shown below.

Person #	1	2	3	4	5	6	7	8	9	10
Height	70	68	60	75	65	69	58	62	66	63

(a) Randomly select five heights from this list by using the random number table that goes with this lesson. Choose a random spot in the table and move down the column. Select the first digit of each number. If you get a repeat, eliminate and keep going. If you get a 0, use this as the 10.

(b) Calculate the **sample mean** to the nearest tenth. Compare to others in the class. What type of variability is being introduced through this process?

average of our 5

$$\frac{68+58+62+66+63}{5}$$

Sample mean = 63.4

Example of sampling varrability.

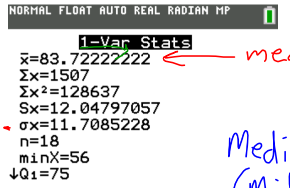
When we conduct a study, the complete set of all subjects that share a common characteristic that is being studied is known as the **population**. All populations have **natural or inter-individual variability**. Most of the time, the entire population is not measured, but a sample is taken to infer characteristics of a population. Still, all populations in theory have **population parameters** that describe the population, such as its mean, standard deviation, and interquartile range.

***Exercise* #2:** 18 students in Mr. Weiler's Advanced Calculus class took a quiz with the following results in ascending order.

Q1

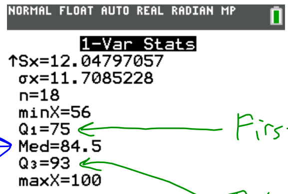
56, 68, 72, 72, 75, 78, 80, 84, 84, 85, 88, 88, 90, 93, 95, 99, 100, 100

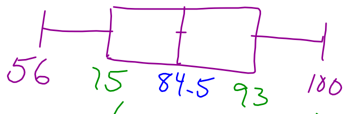
(a) Use your calculator to determine the mean, the median, and the quartiles for this data set. Then, construct a simple box-and-whiskers (box plot) for this data set.



mean

Median (Middle)





First Quartile (Lower quartile)

Third Quartile (Upper Quartile)

(b) What is the interquartile range of this data set? In theory, what percent of the data set should lie between the first and third quartiles? Is that true for this data set?

$$Q_3 - Q_1 = \quad 93 - 75 = \boxed{18}$$

(c) What is the population standard deviation for this data set to the nearest tenth? How do you interpret the standard deviation?

$$\sigma x = 11.7$$

On average a data point is 11.7 units from the mean.

(d) What percent of the scores were within one standard deviation of the mean? Within two standard deviations of the mean? Round your percents to the nearest percent and show your work.

Within One Standard Deviation of the Mean

σx $\quad$ $\bar{x}$

$\bar{x} + \sigma x = 83.7 + 11.7 = 95.4$

$\bar{x} - \sigma x = 83.7 - 11.7 = 72$

$11/18 = 61\%$

Within Two Standard Deviations of the Mean

$2\sigma x$ $\quad$ $\bar{x}$

$\bar{x} + 2\sigma x$

$83.7 + 2(11.7) = 107.1$

$\bar{x} - 2\sigma x$

$83.7 - 2(11.7) = 60.3$

$\frac{17 \text{ scores}}{18} = 94\%$

Sometimes data is grouped in a frequency chart. We still should be able to calculate the basic population parameters when the information is given in this form.

***Exercise* #3:** A small company has salaries for their 50 employees as given in the table below

Salary (x_i) L1	Frequency (f_i) L2
25,000	5
32,000	21
45,000	14
58,000	7
75,000	2
120,000	1

(a) Find the mean and standard deviations of the salary range.

$\bar{x} = 42060$ (mean) $\sigma x = 16567.93$ (std dev)

(b) What is the median of this data set? Why is the median considerably lower than the mean in this data set?

median = 32000

b/c there are 21 people making 32000

There is one person making considerably more (120,000)

(c) Does more or less than 50% of the data set fall within one standard deviation of the mean? Show the analysis that leads to your answer.

$\bar{x} + \sigma x = 42060 + 16567.93 = 58627.93$ 42 out of 50.

84%

$\bar{x} - \sigma x = 42060 - 16567.93 = 25492.07$

Although we have often concentrated on experimental studies where data is collected and means are found, many times we use statistics to represent results of a survey where we are interested in what **proportion** of a **population** share a certain characteristic. These proportions are most expressed as decimals, but sometimes are represented by fractions or percents.

***Exercise* #4:** A questionnaire went home to all juniors concerning their ability to bring and use mobile devices at school. The questionnaires constituted a **census** since all of the juniors were surveyed. Of the 742 juniors, 564 of them reported having web-enabled mobile devices. What was the population proportion for web-enabled devices? Express your answer as a decimal and as a percent.

***Exercise* #5:** The proportion of eggs that get cracked in a local egg handling facility is 0.023. If 2,500 dozen eggs are packaged in the factor per day, what should we expect to be the number of eggs cracked per day?

(1) 350 (3) 230

(2) 450 (4) 690

Normal Curve / Bell Curve

(Applies when data is considered normal)